

## JP9330316A: TRANSLATING METHOD FOR ELECTRONIC MAIL

[View Images \(1 pages\)](#)

Inventor(s): **TAKAHASHI KAZUHIKO**

Applicant(s): **MEIDENSHA CORP**

Issued/Filed Dates: **Dec. 22, 1997 / June 10, 1996**

Application Number: **JP1996000146848**

IPC Class: **G06F 017/27; G06F 013/00; G06F 017/28; G06F 017/21;**

Abstract: **Problem to be solved:** To provide an electronic mail translating method which is improved in translation precision and made easy in check operation for transmission, by converting a recognized document into the language of the transmission country by inference using the inference rule of a knowledgebase.  
**Solution:** Word recognition (S31, S32) of an inputted document using the word dictionary of the knowledge base, syntax analysis (S33, S34) using the grammatical dictionary of the knowledge base, and meaning analysis (S35, S36) using the meaning dictionary of the knowledge base are performed by input(S1) of the document and selection (S2) of the language of the transmission country, the recognized document is converted (S37, S38, S39) into the language of the transmission country by inference using the inference rule of the knowledge base, and the converted document is generated and checked (S310, S4) and transmitted to the transmission country through the Internet under communication control (S5-S10).  
 COPYRIGHT: (C)1997,JPO

Foreign References: **none**

(No patents reference this one)



**Nominate this invention for the Gallery...**

**Alternate Searches**



[Patent Number](#)



[Boolean Text](#)



[Advanced Text](#)



**SEARCH PATENT FULL TEXT WITH NATURAL LANGUAGE**

-22- (WPAT)

AN - 98-106746/10

KRPX- N98-085727

TI - Language translation method for E-mail system - involves translating language of document whose native language is recognised by performing syntactic/implication analysis into language of destination country based on reasoning rule

DC - T01

AW - INTERNET

PA - (MEID ) MEIDENSHA CORP

PR - 96.06.10 96JP-146848

NUM - 1 patent(s) 1 country(s)

PN -- JP09330316 A 97.12.22 \* (9810) 8p G06F-017/27

AP -- 96JP-146848 96.06.10

IC1 - G06F-017/27

IC2 - G06F-013/00 G06F-017/21 G06F-017/28

AB - JP09330316 A

The method involves using a translation unit which recognises the words contained in an input document using a word dictionary. The language of the input document is recognised, by performing syntactic analysis using a grammar dictionary and implication analysis using an implication dictionary.

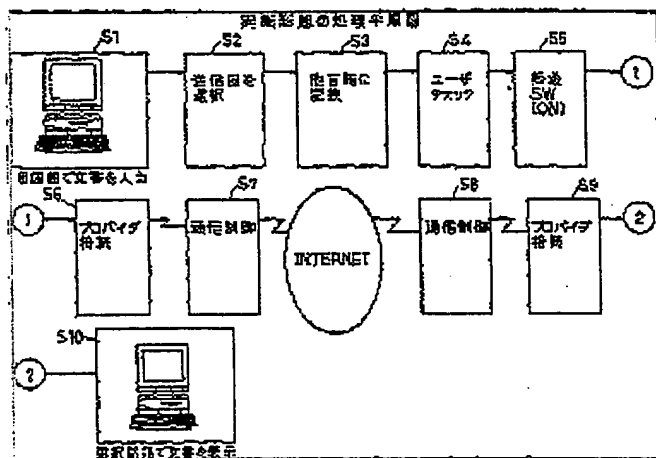
The document is converted into language of a destination country based on a reasoning rule. The converted document is transmitted to the destination using internet.

ADVANTAGE - Facilitates wide translation range. Improves translation accuracy.

(Dwg.2/10)

MC - EPI: T01-J11A T01-J14

FN - WPI2AD61.GIF



(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-330316

(43) 公開日 平成9年(1997)12月22日

(51) Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 17/27			G 0 6 F 15/38	D
13/00	3 5 1		13/00	3 5 1 G
17/28			15/20	5 9 2 A
17/21				5 9 6 A

審査請求 未請求 請求項の数 2 O L (全 8 頁)

(21) 出願番号 特願平8-146848

(22) 出願日 平成8年(1996)6月10日

(71) 出願人 000006105

株式会社明電舎

東京都品川区大崎2丁目1番17号

(72) 発明者 高橋 和彦

東京都品川区大崎2丁目1番17号 株式会  
社明電舎内

(74) 代理人 弁理士 志賀 富士弥 (外1名)

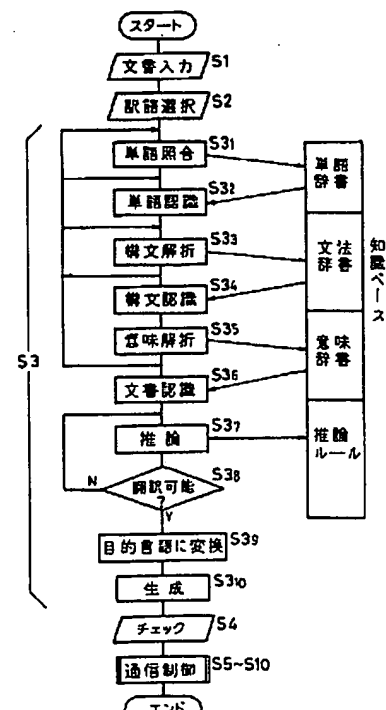
(54) 【発明の名称】 電子メールの翻訳方法

## (57) 【要約】

【課題】 従来の電子メールの翻訳方式では、翻訳処理のための手直し作業が多く、翻訳精度的にも不十分なものであった。さらに、電子メールシステムとは分離されるため、システムとの融合性に欠ける。

【解決手段】 文書の入力 (S1) と送信国の言語の選択 (S2) により、入力された文書を知識ベースの単語辞書を使った単語認識 (S3<sub>1</sub>、S3<sub>2</sub>) と、知識ベースの文法辞書を使った構文解析 (S3<sub>3</sub>、S3<sub>4</sub>) 及び知識ベースの意味辞書を使った意味解析 (S3<sub>5</sub>、S3<sub>6</sub>) を行うことで母国語の文書認識を行い、この認識した文書について知識ベースの推論ルールを使った推論で送信国の言語に変換し (S3<sub>7</sub>、S3<sub>8</sub>、S3<sub>9</sub>)、この変換した文書を生成・チェックし (S3<sub>10</sub>、S4)、通信制御によりインターネットを通して送信国に送信する (S5～S10)。

他言語への変換処理手順図



## 【特許請求の範囲】

【請求項1】 母国語の文書を翻訳システムを使って送信国の言語の文書に変換し、この変換した文書を送信国ヘインターネットで送信する電子メールシステムにおいて、

前記翻訳システムは、入力された文書について、知識ベースの単語辞書を使った単語認識と、知識ベースの文法辞書を使った構文解析及び知識ベースの意味辞書を使った意味解析を行うことで母国語の文書認識を行い、この認識した文書について知識ベースの推論ルールを使った推論で送信国の言語に変換することを特徴とする電子メールの翻訳方法。

【請求項2】 母国語の文書を翻訳システムを使って送信国の言語の文書に変換し、この変換した文書を送信国ヘインターネットで送信する電子メールシステムにおいて、

前記翻訳システムは、入力された文書について、ファジイ推論により単語及び文の曖昧性を除去した文を得、この文に対して知識ベースの文法辞書を使った構文解析及び知識ベースの意味辞書を使った意味解析を行うことで母国語の文書認識を行い、この認識した文書について送信国の言語に変換することを特徴とする電子メールの翻訳方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、母国語を翻訳ソフトで外国語に変換し、インターネットを利用して海外へ文書を電子（E）メールで送信するための電子メールシステムに係り、特に送信文書を母国語から外国語に翻訳するための電子メールの翻訳方法に関する。

## 【0002】

【従来の技術】電子メールによる文書の送信において、母国語を送信先の外国語に自動変換する方式には、トランスファー方式、ダイレクト方式、ピボット方式、さらにはこれらの方式の特徴を組み合わせた方式がある。

【0003】トランスファー方式は、言語別の中間表現を持ち、中間表現での変換が行われる。例えば、英語→日本語、仏語→日本語などを翻訳する言語に中間表現がそれぞれ生成される。また、解析手法としては日本語には格文法、英語には拡張遷移文法がよく利用される。

【0004】図8は、トランスファー方式による翻訳システム構成を例文と共に示す。元の言語Aとしての日本語の文「明日は、晴れでしょう」を目的の言語Bとしての英語の文に変換するのに、元の文を単語辞書C及び文法辞書Dを使った日本語処理方法による構文解析と意味解析を行う構文・意味解析Eを行うことで文節「明日は」と「でしょう」と「晴れ」の組み合わせになる中間表現Fを得る。

【0005】この中間表現Fを「tomorrow」と「It will be fine」の組み合わせにする

英語の中間表現Gに変換する。この中間表現Gから辞書C、Dを使った英語の表層文生成Hにより目的の言語B「It will finetomorrow」を生成する。

【0006】図9は、トランスファー方式とダイレクト方式及びピボット方式の3つの方式を組み合わせた変換方式を示す。トランスファー方式は、前記のように元の言語Aと変換目的の言語Bにそれぞれ特徴的な中間表現F、Gを生成した変換を行う。

10 【0007】ダイレクト方式は、元の言語の単語を目的の言語の単語に置き換え、その単語の順序を並べ換えて変換し、ピボット方式は、元の言語の種類に依存しない共通の中間表現Iに変換し、その中間表現から目的とする言語を生成するが、これら両方式は、トランスファー方式に比べて変換精度が劣る。

【0008】なお、上記の3つの変換方式の特徴を合わせ、より複雑な翻訳処理を可能にする方式もある。

20 【0009】また、市販されているソフトの類は、上記の何れかの方法の対訳エディタ、ユーザ辞書登録、訳語学習、翻訳制御、パターン翻訳等のオンライン翻訳機能を備えており、この翻訳システムの特徴として、例えば、

（1）基本辞書（英日8万語、日英9万語）と豊富な訳語で表現力豊かな翻訳結果が得られる。

【0010】（2）34分野130万語の専門用語辞書が使用分野の翻訳をカバーする。

【0011】（3）ユーザインタフェースの導入により、特殊な操作や不必要な機能は省き、翻訳作業を軽減させる。

30 【0012】といったものがある。

## 【0013】

【発明が解決しようとする課題】従来の翻訳システムにおける処理過程は、図10に示すようになり、母国語になる原文をコンピュータが読解かつ翻訳可能な文に訂正するプレエディット作業により電子メールの入力文を得、これを翻訳システムにより翻訳して外国語の出力文を得、この出力文をユーザがスペルミスや文法ミスをチェック・訂正するポストエディット作業を行い、電子メールの送信文とする。

40 【0014】このため、翻訳処理のための手直し作業が多く、文書を相手へ送信するまでに多くの時間と作業が必要とする。

【0015】特に、曖昧性を持つ文書には翻訳ミスが多くなり、これを修正するためのプレエディット作業に手間取る。

【0016】また、従来の翻訳システムは、翻訳語の範囲（翻訳可能な言語数、専門用語の範囲）が限られ、翻訳精度的に不十分なものであった。

【0017】さらに、従来の翻訳システムは、電子メールシステムとは分離されたソフトになり、電子メールシ

システムとの融合性に欠け、使い勝手を悪くしている。

【0018】本発明の目的は、翻訳語の範囲を広げながら翻訳精度を高め、送信のためのチェック作業を簡単にした電子メールの翻訳方法を提供することにある。

【0019】本発明の他の目的は、電子メールシステムとの融合性を高めた電子メールの翻訳方法を提供することにある。

【0020】

【課題を解決するための手段】本発明は、母国語の文書を翻訳システムを使って送信国の言語の文書に変換し、この変換した文書を送信国へインターネットで送信する電子メールシステムにおいて、前記翻訳システムは、入力された文書について、知識ベースの単語辞書を使った単語認識と、知識ベースの文法辞書を使った構文解析及び知識ベースの意味辞書を使った意味解析を行うことで母国語の文書認識を行い、この認識した文書について知識ベースの推論ルールを使った推論で送信国の言語に変換することを特徴とする。

【0021】また、本発明は、母国語の文書を翻訳システムを使って送信国の言語の文書に変換し、この変換した文書を送信国へインターネットで送信する電子メールシステムにおいて、前記翻訳システムは、入力された文書について、ファジィ推論により単語及び文の曖昧性を除去した文を得、この文に対して知識ベースの文法辞書を使った構文解析及び知識ベースの意味辞書を使った意味解析を行うことで母国語の文書認識を行い、この認識した文書について送信国の言語に変換することを特徴とする。

【0022】

【発明の実施の形態】

(第1の実施形態) 図1は、本発明の実施形態を示す電子メールシステムの処理手順である。

【0023】まず、送信側端末になるコンピュータの画面上で送信しようとする文書を母国語で入力する(S1)。次に、入力された文書の送信国をコンピュータの画面上で選択する(S2)。

【0024】この送信国の選択は、送信側の画面例を図2に示すように、送信国名をダイアログボックス表示し、そのうちの1つを選択することになる。

【0025】次に、処理S1で入力された文書について、選択された送信国の言語に変換する(S3)。この変換された文書をユーザがチェックし(S4)、チェック終了で転送スイッチをオン(例えば、図2の画面上の転送ボタンGUIを選択)操作する(S5)。

【0026】この変換された文書の転送開始に際して、コンピュータは通信回線を介してプロバイダと接続して送信要求を行う(S6)。プロバイダは、送信要求で所定のプロトコルによる通信制御(S7)とインターネットを介した送信国側の通信制御(S8)で送信国側のプロバイダと接続を試みる(S9)。接続に成功したとき、送信文書の送信を行う。この文書は送信国側のプロバイダを介して送信国側の受信端末になるコンピュータで受信される(S10)。

送信文書の送信を行う。この文書は送信国側のプロバイダを介して送信国側の受信端末になるコンピュータで受信される(S10)。

【0027】以上までの処理により、本実施形態では、翻訳システムと電子メールシステムを融合させ、文書を送信する際にシステム上の選択送信国の中から送信国名を選択することで送信文書を送信国の言語に変換し、ユーザのチェックと転送操作のみでインターネットを介した送信ができる。

【0028】なお、変換した文書の送信方法は、直接にプロバイダに接続するに限らず、送信端末や受信端末がLANやWANで接続される方式のもの、ISDNによるデジタル通信によるものなど通信方式は他の方式のものでも良い。

【0029】以上までの処理手順のうち、母国語から外国語への変換処理の詳細を図3に示す。同図では、図1の処理S1～S10と対応付けて示し、処理S3の部分に変換処理の詳細である。

【0030】言語変換処理S3は、知識ベースに用意する単語辞書と文法辞書及び意味辞書及び推論ルールを参照して目的言語に変換するものである。この変換には、まず、入力された母国語の文書に含まれる単語について学習機能を持つ単語辞書との照合(S3<sub>1</sub>)によって個々の単語を認識する(S3<sub>2</sub>)。

【0031】次に、学習機能を持つ文法辞書を参照した文書の母国語の統語規則から文書の構文解析を行い(S3<sub>3</sub>)、文書の構文を認識する(S3<sub>4</sub>)。この認識は、単語の係り受けなど文法的に適合した構文木(文のパターン)を得るもので、1つの文書について複数の構文木が作成される場合がある。

【0032】この構文的な曖昧性を持つ複数の構文木から正しい構文を選択するために、学習機能を持つ意味辞書の参照で文書の意味的な情報を取得し、意味的に不適当な構文木を削除する意味解析を行い(S3<sub>5</sub>)、文書の意味認識を行う(S3<sub>6</sub>)。

【0033】上記の構文解析及び意味解析において、単語の照合や認識、さらには構文解析を再度行うこともある。

【0034】次に、単語と構文と意味を認識した文書について、学習機能を持つ知識ベースの推論ルールを使ったエキスパートシステムになる推論で目的言語への翻訳を行う(S3<sub>7</sub>)。この推論で翻訳が可能となったとき(S3<sub>8</sub>)、翻訳処理を実行して目的言語に変換し(S3<sub>9</sub>)、送信文を生成する(S3<sub>10</sub>)。

【0035】以上のように、本実施形態では、母国語を外国語に変換するのに、知識ベースとして持つ辞書を使って母国語の単語、構文、意味の認識・解析を行う自然言語処理を行い、この処理を終えた文書をエキスパートシステムを使った推論で外国語に変換する。これにより、中間表現の変換を行ってコンパイル、変換処理を行う。

て変換制度を高めた推論による変換になり、変換前及び変換後の手直し作業を減らすことができる。

【0036】また、知識ベースを持たせることにより、訳語や専門用語の範囲を広げ、言語間の選択範囲を広げることができる。

【0037】また、知識ベースは、各種の辞書の参照に学習機能を持たせることにより、自然言語処理及び推論による変換精度・効率を一層高めることができる。

【0038】(第2の実施形態)図4は、本発明の他の実施形態を示し、図1における他言語への変換処理S3 10の部分である。

【0039】本実施形態では、言語変換処理S3は、入力文書の文体解析を行い(S31)、ファジィ推論により曖昧な語彙を含む入力文書からコンピュータが解読可能な文を得る(S32)。このファジィ推論は、曖昧な語彙を省略し、ルールによっては単語を知識ベースの単語辞書へ照合し(S33)、必要な単語文を削除あるいは抽出して単語を認識する(S34)。

【0040】この後、知識ベースに用意する文法辞書及び意味辞書による自然言語解析により構文解析(S35, S36)及び意味解析(S37, S38)を行う。

【0041】次に、単語と構文と意味を認識した文書について、翻訳辞書を使った翻訳変換のプロセスに入り、翻訳不可能ならば(S39)再度、ファジィ推論による文の解析に戻り(S32)、ファジィ推論によって翻訳可能な文にまとめ、翻訳を実行する(S310)。

【0042】また、翻訳処理の実行中は、選択訳語の単語、文法、構文を知識ベースに照合し、訳語を序列化する(S311)。

【0043】以上までの変換処理において、ファジィ推 30論により曖昧な原文からコンピュータが解読可能な文を\*

U = (1・1・-0.9・0.1・0.8・0.8・-0.1・-0.9)						
宣	到	「	住	島	2	ひ
教	若	、	ん	に	っ	と
師	し	」	で	の	り	」
が	ま		い	部	で	
	し		る	族		
	た			が		

【0052】とし、これをマムダニの方法を例にとり、図5に示すように、ファジィラベルとしてPB(positive big), PM(positive medium), PS(positive small), ZO(zero), NS(negative small)を用い、さらに様々なファジィ変数に共通して使えるように、-6以上6以下に規格化した台集合上で規定する。

【0053】また、上記を三角形の連続関数として表す 50図6に示すようになる。さらに max-minに\*

\*得る処理を以下に具体的に説明する。

【0044】入力文「2つの部族が住んでいる島に、ひとりの宣教師が到着しました。」を解読対象とすると、以下の3つのルールの中から文を解読するのに必要とされるルールを1つ選択し、処理を実行する。

【0045】(a)条件「コンピュータが解読可能ならば」、結論「コンピュータが認識」。

【0046】(b)条件「コンピュータが解読不可能であれば」、結論「コンピュータが一部の文を削除」。

【0047】(c)条件「コンピュータが解読不可能であれば」、結論「コンピュータが知識ベースへ照合」。

【0048】上記のルール(a)の場合、入力文をコンピュータが解読可能ならば、コンピュータが文の認識を行い、解釈不可能ならば文の解読を行う。ルール(b)の場合、max-min法でコンピュータが解読できる高い数値にあるならば解読できる文として採用し、低い数値のものは削除する。ルール(c)は、あらかじめコンピュータに認識されていないため解読不可能な文又は語彙を学習機能を持つ単語辞書に対する単語照合を実行 20する。

【0049】次に、制御でよく使用されるマムダニの方法を例に取りファジィ推論を説明する。制御状態(e, Δe)をルール(a), (b), (c)に代入し、それぞれの出力ΔU1, ΔU2, ΔU3が推論結果として得られ、重ね合わせた総面積の重心が、結果ΔUとして出力される。

【0050】上記の入力文では、まず最初に行列形式で、

【0051】

【表1】

※よって文の語彙を除去する領域を特定すると、図7に示すように表示される。

【0054】同図の0.1〜0.1の領域は、コンピュータがその文を採用か、不採用かの判断がつかねる場合、ルール(c)として知識ベースへ照合する「知識ベースゾーン」とする。その結果、入力文はコンピュータが解読可能な文として、「宣教師が2つの部族が住む島に到着しました」となる。

【0055】以上のように、本実施形態によれば、曖昧な文書の入力にはファジィ推論により曖昧な単語・文節

を削除して文書の翻訳に適した文に変換し、これを目的言語に翻訳することにより、目的言語への変換精度を一層高めることができ、これに伴いユーザのチェック作業も簡単になる。

【0056】なお、認識した文書を目的言語に変換するのに、エキスパートシステムの推論ルールによる変換とすることでも良い。

【0057】

【発明の効果】以上のとおり、本発明によれば、母国語の文書を翻訳システムを使って送信国の言語の文書に変換するのに、入力された文書について、知識ベースの各辞書を使った単語認識と、構文解析及び意味解析を行うことで母国語の文書認識を行い、この認識した文書について知識ベースの推論ルールを使った推論で送信国の言語に変換するようにしたため、以下の効果がある。

【0058】(1) 知識ベースを使った文書認識と推論により翻訳できる言語の範囲を広げることができる。

【0059】(2) 文書認識と推論により翻訳精度を高めることができる。

【0060】(3) 翻訳精度を高めることで、ユーザのチェック作業が簡単になる。

【0061】(4) 送信文書の入力と送信国の言語の選

択及び簡単なチェックにより送信国の言語による送信ができ、電子メールシステムとの融合性を高めることができる。

【0062】また、本発明によれば、入力文書をファジイ推論により曖昧性を除去した文に変換するようにしたため、曖昧性を持つ文書の入力にも翻訳精度を高めることができる。

【図面の簡単な説明】

【図1】本発明の実施形態を示す電子メールシステムの処理手順図。

【図2】実施形態における送信側の画面例。

【図3】実施形態における他言語への変換処理手順図。

【図4】本発明の他の実施形態を示す他言語への変換処理手順図。

【図5】他の実施形態におけるファジイ推論による文書解読の推論例。

【図6】他の実施形態におけるメンバーシップ関数例。

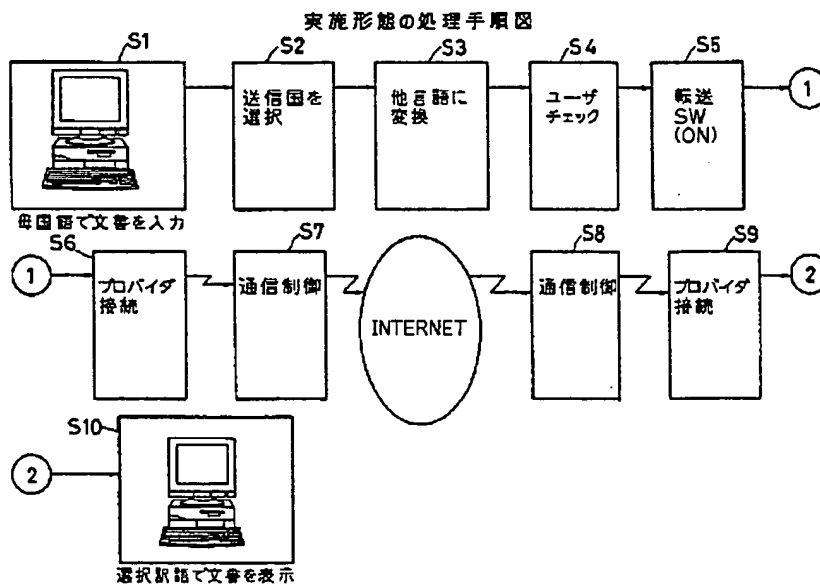
【図7】他の実施形態における語彙を除去する領域例。

【図8】従来のトランスファー方式による翻訳システム構成図。

【図9】従来の3つの変換方式。

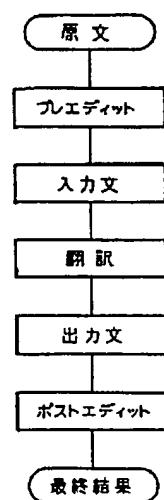
【図10】従来の翻訳の処理過程図。

【図1】



【図10】

従来の翻訳の処理過程



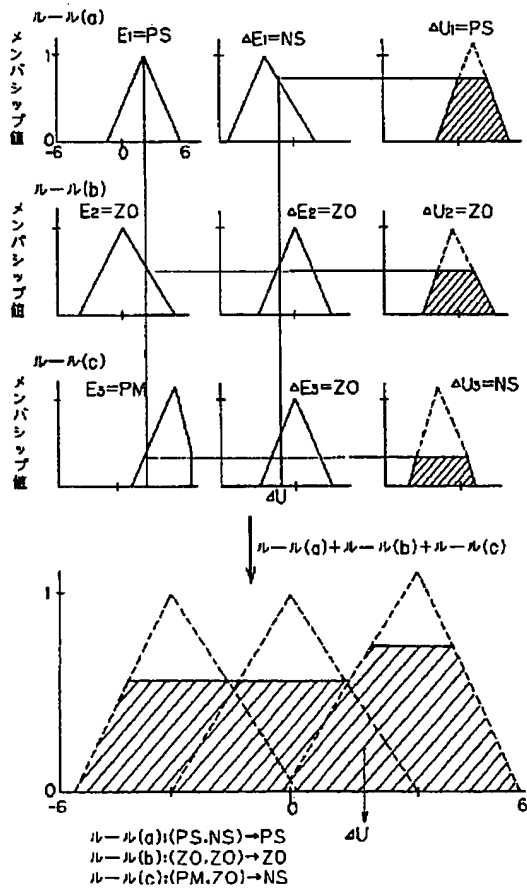
【図2】

送信側の画面例

マシン名		送信国	1. 米国
ユーザネーム			2. 英国
パスワード			3. 独 国
送信する文書名			4. 仏 国
			5. 中 国
			6. 韓 国
			7. 伊 国
転送	ON		

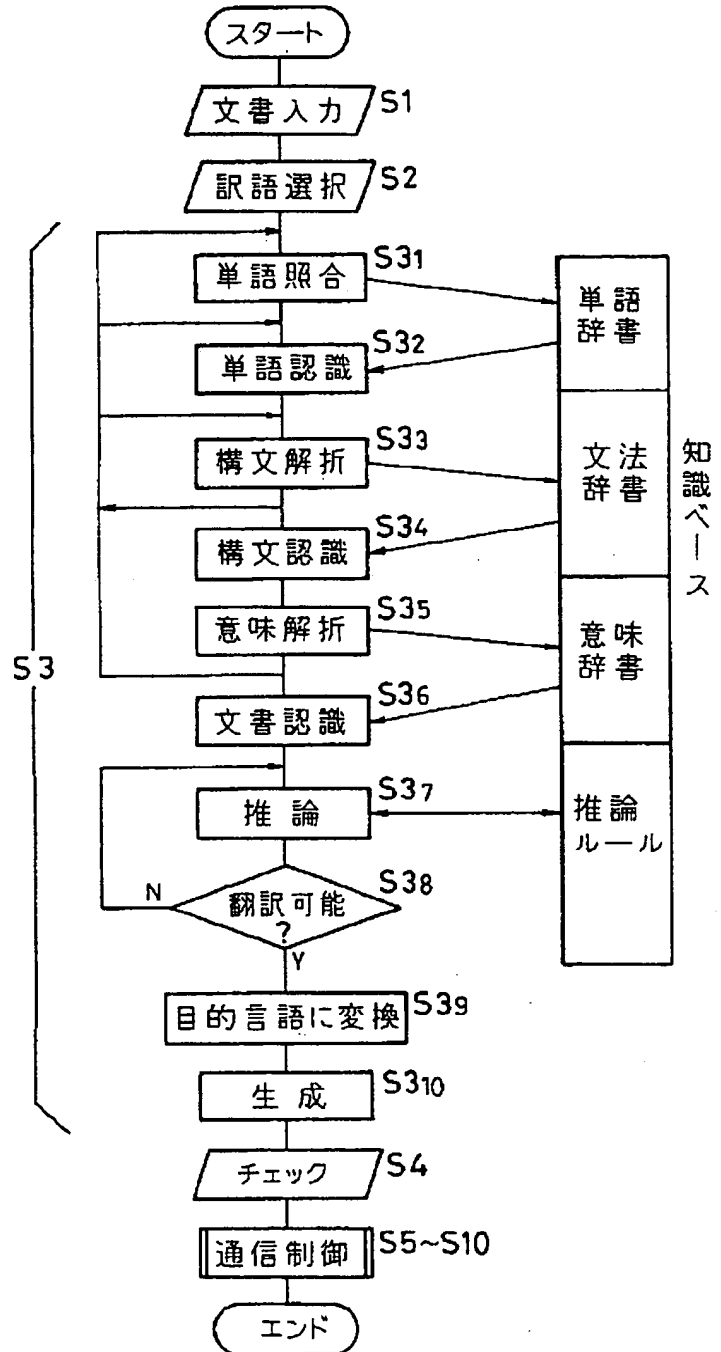
【図5】

文書解読の推論



【図3】

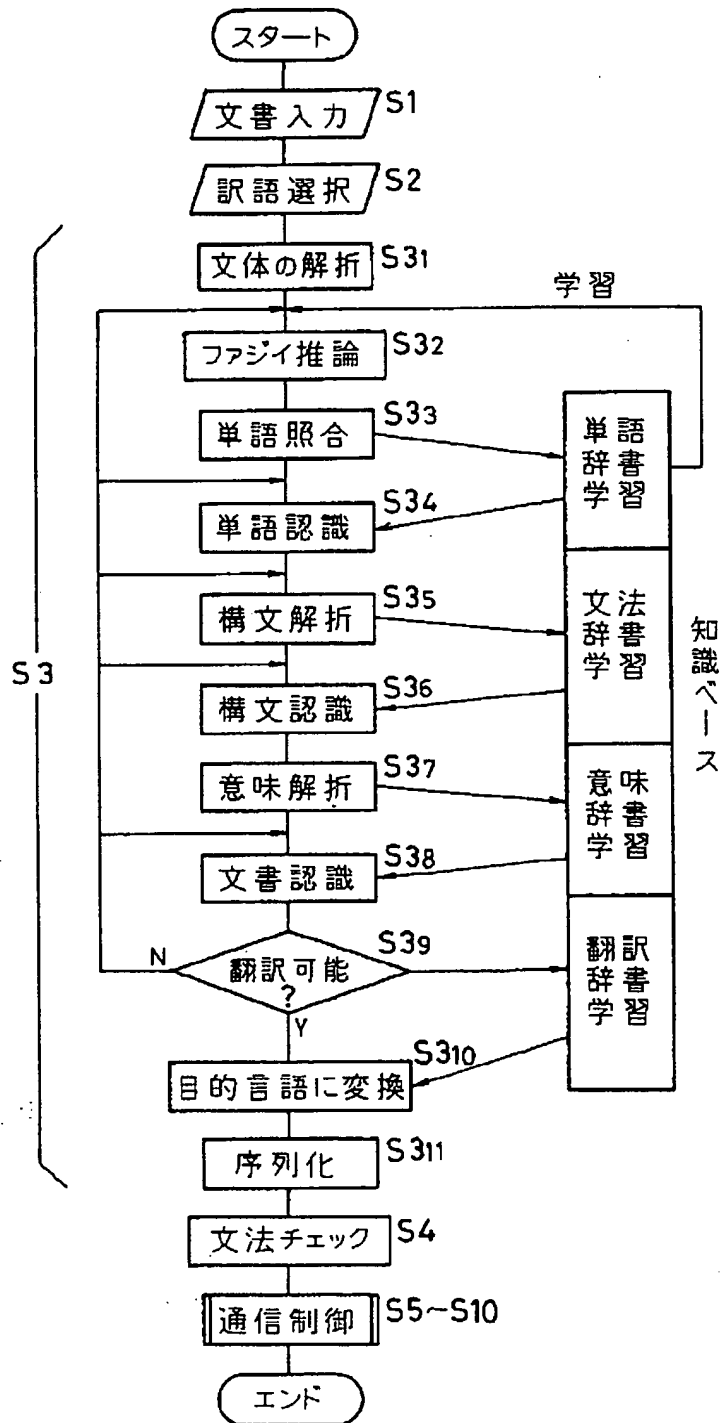
他言語への変換処理手順図





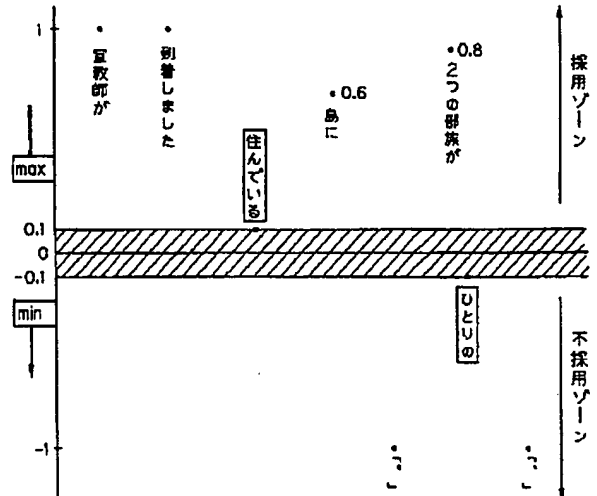
【図4】

他言語への変換処理手順図



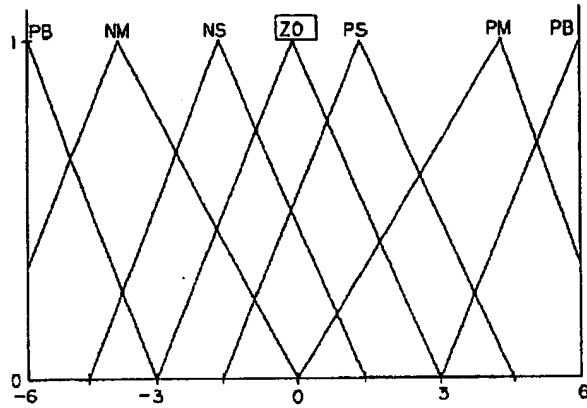
【図7】

語彙を除去する領域



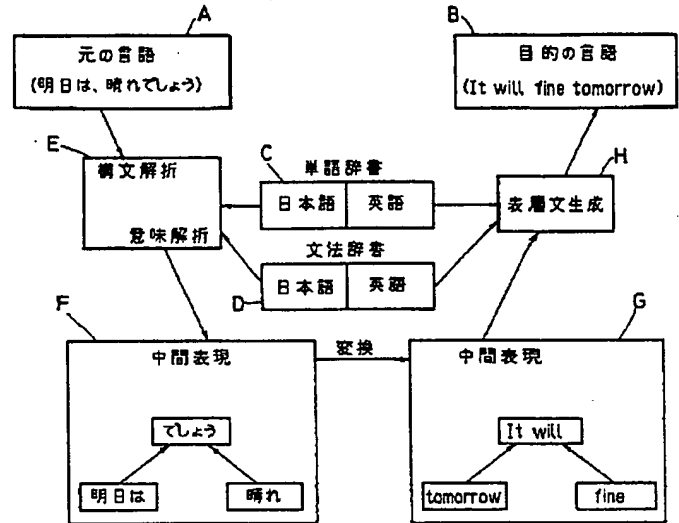
【図6】

メンバーシップ関数



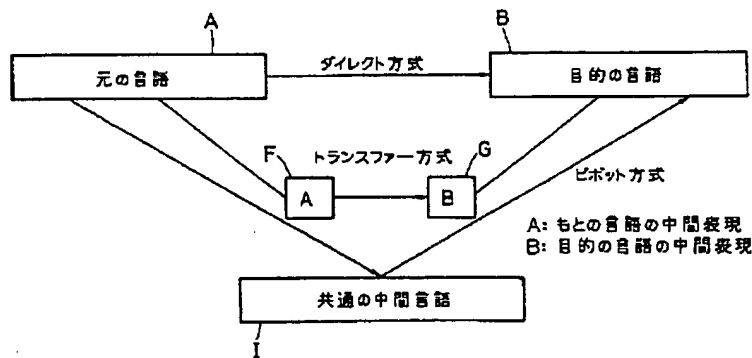
【図8】

トランスファー方式による翻訳システム構成図



【図9】

3つの変換方式



mention  
in prior  
art

This is a translation method that relies on a knowledge base. It could be used in the Worldhingo patent. Translation occurs before transmit

- (19) Patent Office of Japan (JP)
- (12) Public Advertisement on Publicised Patent (A)
- (11) Patent Application Publicity Reference Number: Toku-Kai-Hei 9-330316
- (43) Date of Publicity: 22 December 1997

(51) Int.Cl. <sup>8</sup>		Distin- guisher	Internal Ref.	F1		Technological Reference
G06F	17/27			G06F	15/38	D
	13/00	351			13/00	351G
	17/28				15/20	592A
	17/21					596A

Examination applied    Unapplied    No. of applications: 2    OL    (No. of pages: 8)

- (21) Application Number: Toku-Kai-Hei 8-146848
- (22) Date of Application: 10 June 1996

- (71) Name of Applicant: 000006105 Meidensha Co. Ltd., 2-1-17 Osaki, Shinagawa-ku, Tokyo
- (72) Name of Inventor: Kazuhiko Takahashi, c/o Meidensha Co. Ltd., 2-1-17 Osaki, Shinagawa-ku, Tokyo
- (74) Representative: Patent Attorney Fujiya Shiga, et al.

(54) [Title of Invention] Method of Translating Electronic Mails

(57) [Summary]

[Problem] Conventional methods of translating electronic mails required substantial editing before and after processing as their accuracy was insufficient. Furthermore, as they were separated programmes, they lacked integrity with the [(e-mail)] system.

[Solution] Firstly, an operator inputs the document (S1) and selects the language of destination (S2). The document is recognised through the following sequence: vocabulary recognition using the vocabulary dictionary component of the "Knowledge Base" (S3<sub>1</sub>, S3<sub>2</sub>); syntax recognition using the grammar dictionary component of the "Knowledge Base" (S3<sub>3</sub>, S3<sub>4</sub>); and semantic analysis using the semantic dictionary component of the "Knowledge Base" (S3<sub>5</sub>,

S3<sub>6</sub>). The recognised document is then translated using the inference rules component of the "Knowledge Base" into the language of destination (S3<sub>7</sub>, S3<sub>8</sub>, S3<sub>9</sub>). The document thus translated is generated and checked (S3<sub>10</sub>, S4) before being transmitted to the destination via internet through communication administration system (S5 – S10).

[Scope of the Patent Application]

[Application Item 1] In the electronic mail system that translates the original documents into the language of the destination through a translation system ~~before transmitting the translated documents to the destination via internet.~~

This translation system has a unique method of translating the electronic mail into the language of the destination. Firstly it conducts original language sentence recognition using the vocabulary dictionary component, the grammar dictionary component and the semantic dictionary component of the "Knowledge Base". The document thus recognised is then translated into the language of the destination using knowledge-based inference rules.

[Application Item 2]

In the electronic mail system that translates the original documents into the language of the destination through a translation system ~~before transmitting the translated documents to the destination via internet.~~

This translation system applies fuzzy-inference to the input document in order to discard ambiguities from its vocabulary and syntax. Sentences thus extracted are then recognised through a syntax analysis using the grammar dictionary component of the "Knowledge Base" and a semantic analysis using the semantic dictionary component of the "Knowledge Base". The document thus recognised is then translated into the language of the destination.

[Detailed description of the invention]

[0001]

[Technological categorisation of the invention] This invention concerns the electronic (E) mail system by which electronic mails written in the original language are translated into foreign languages by a translation software before being sent overseas through the internet. It also concerns a method of translating electronic mails from the original language into foreign languages.

[0002]

[Conventional technology] Methods of automatic translation for translating electronic mails from the original language into foreign languages include the transfer method, the direct method, the pivot method and the combination of these methods.

[0003] The transfer method involves intermediary for each language: the translation is done by using two intermediaries. For example, when translating from English to Japanese, or from French to Japanese, a unique intermediary is produced in each process. As to the methods of sentence analysis, case-oriented analysis is often used for Japanese. In the case of English, inflectional morphological grammar is often used.

[0004] Figure 8 describes, by an example, the structure of a translation system using the transfer method. The original language A in this example is Japanese. A sentence in language A, "Asu wa hare deshou" is being translated into the destination language B which, in this case, is English. The original sentence is processed by a Japanese syntax analyser using the vocabulary dictionary (C) and the grammar dictionary (D). The sentence is then processed by a syntax-semantics analyser (E), which produces an intermediary (F). The intermediary is a combination of "asu wa", "deshou" and "hare".

[0005] The intermediary (F) is then translated into an English intermediary (G), which is a combination of "tomorrow", "it will" and "fine". This intermediary (G) is then processed by expression generator (H) using dictionaries (C) and (D). The sentence thus generated in destination language B reads "It will finetomorrow [sic]".

[0006] Figure 9 describes translation method using transfer, direct and pivot methods in combination. As described above, transfer method generates unique intermediaries F and G other than the original A and the destination language B.

[0007] The direct method translation translates the original text word by word and then shuffles the translated words. Pivot method generates only one intermediary (I) irrespective of the language of the original before translating it into the destination language. Both are inferior to the transfer method in terms of accuracy.

[0008] There are methods that combine the three methods discussed above in order to enable complex translations.

[0009] The translation softwares currently available are equipped with an interrelational text editor based on one or more translation methods described above, user-programmable dictionaries, word-by-word equivalents memoriser, translation administrator, pattern translator and so on. They are equipped with on-line translation functions. The features of this translation system include:

(1) The main dictionary that contains 80,000 English-to-Japanese translations and 90,000 Japanese-to-English translations. The ample vocabulary enables it to produce colourful translation results.

[0010] (2) Technical dictionaries list 1,300,000 words categorised in 34 subjects, catering users' specialty fields.

[0011] (3) A user interface designed to reduce complicated operations and unnecessary functions simplifies the operation when translating.

[0012] These are some of the features of this system.

[0013]

[The Problems this invention attempts to address] Figure 10 describes the process of translation by conventional systems. Pre-editing of the original document is required in order to make it recognisable and translatable by the computer. The pre-edited sentences thus obtained are then translated into a foreign language by the system. The translated sentences are then checked by the user to correct any spelling or grammatical errors (post-editing). The sentences thus corrected are used in the electronic mail to be transmitted.

[0014] As described above, conventional translation systems required substantial manual works. As a result, much time and effort are required before a document can be transmitted to the destination.

[0015] Ambiguous sentences are particularly susceptible to translation errors. Consequently, pre-editing of the document in order to eliminate ambiguities may take much effort.

[0016] Furthermore, conventional translation systems are limited in terms of vocabulary (number of languages that can be translated and/or subject areas of specialised terms), thus leaving the accuracy of translation results unsatisfactory.

[0017] Furthermore, conventional translations systems are separate from the electronic mail systems. As such, they lack integrity with the electronic mail systems and are cumbersome to use.

[0018] The objective of this invention is to increase the number of translatable languages while at the same time improving the accuracy of translation, in order to reduce the necessity of checks by the users before transmission.

[0019] Another objective of this invention is to provide the users more integrated electronic mail and electronic mail translation systems.

[0020]

[Measures to solve the problem] This invention is a translation system within an electronic mail system that translates a document the user wrote in his or her own language into the language of the destination before transmitting it via internet. The translations are conducted as follows: firstly it conducts original language sentence recognition using the vocabulary dictionary component, the grammar dictionary component and the semantic dictionary

component of the "Knowledge Base"; the document thus recognised is then translated into the language of the destination using knowledge-based inference rules.

[0021] This invention is a translation system within an electronic mail system that translates a document the user wrote in his or her own language into the language of the destination before transmitting it via internet. This translation system applies fuzzy-inference to the input document in order to discard ambiguities from its vocabulary and syntax. Sentences thus extracted are then recognised through a syntax analysis using the grammar dictionary component of the "Knowledge Base" and a semantic analysis using the semantic dictionary component of the "Knowledge Base". The document thus recognised is then translated into the language of the destination.

[0022]

[Practical application of this invention]

(Practical application: Example 1) Figure 1 illustrates a process of implementing this invention in an electronic mail system.

[0023] Firstly, the user uses his or her terminal to input a document in his or her own language (S1). The user then selects the destination (S2).

[0024] The selection of the destination is done by a dialog box displayed on the user's screen (sample screen image is given in Figure 2). The user selects the country of destination from the list displayed on the screen.

[0025] Secondly, the document inputted at S1 above is translated into the language of the destination (S3). The user then checks the translated document (S4). When the user finishes checking, he or she clicks on a "transmit button" (the "ON" GUI button illustrated in Figure 2) to send the document to the destination (S5).

[0026] Before starting transmission of the document, the computer connects to the provider to request transmission (S6). The provider, upon receiving such request, attempts to establish a connection with another provider in the destination country (S9). The provider regulates the communication via internet by certain protocols (S7) as does the destination's provider (S8). When the communication is successfully established, the document is transmitted. The document is received by the destination's terminal through the destination's provider (S10).

[0027] In this way, this invention achieves integration of the translation and electronic mail systems. The user may simply select the destination from a list of countries provided by the system when sending the document to have the document automatically translated into the language of the destination. The user then needs only to check the translated document and click the transmit button to communicate via internet.

[0028] Incidentally, the translated document may be transmitted by means other than engaging providers. The alternative means include connecting the sender's and receiver's terminals by LAN or WAN, via ISDN or digital network or any other means of communication.

[0029] Figure 3 describes the translation process within the aforementioned system. In relation to Figure 1, the process forms the step S3 of steps S1 –S10 described therein.

[0030] The step S3 above is a translation process using the vocabulary dictionary, grammar dictionary, semantic dictionary and the inference rules provided in the "Knowledge Base". The first procedure of the step is to refer the words in the original document to the vocabulary dictionary, which is capable of learning past references, (S3<sub>1</sub>) and recognise them individually (S3<sub>2</sub>).

[0031] Secondly, it analyses the grammar by referring to the word patterns to the grammar dictionary, which is capable of learning past references, (S3<sub>3</sub>) and recognise the syntax (S3<sub>4</sub>). This recognition yields sentence patterns based on the grammatical rules. There may be several patterns possible for one sentence.

[0032] For a grammatically ambiguous sentence, it is necessary to select one from the several possible sentence patterns. In order to do this, a semantics dictionary is called for to provide semantic information for the document. According to that information, the process eliminates semantically unsuitable sentence patterns (S3<sub>5</sub>) and recognise the document semantically (S3<sub>6</sub>).

[0033] In the process described above, it may become necessary to rerun the vocabulary reference, vocabulary recognition and/or syntax analysis over again.

[0034] The sentence thus recognised in terms of its vocabulary, syntax and semantics, is then translated by an expert system using inference rules of the knowledge base, which is capable of learning past references (S3<sub>7</sub>). If the inference rules are capable of translating the document, the document is translated into the destination's language (S3<sub>8</sub>), and a document for transmission is generated (S3<sub>9</sub>).

[0035] The foregoing example is a translation process from the original language to the destination's language. The process involves natural language recognition, i.e. vocabulary, syntax and semantics recognition and analysis using dictionaries in the "Knowledge Base". An expert system translates the recognised document by inference. The inference translation used in this process yields better accuracy than the intermediary-based transfer and other conventional methods. This advantage reduces the need of manual editing works both before and after the translation.



[0036] Another advantage of this knowledge-based method is that it is possible to extend vocabulary and specialised term dictionaries. At the same time, the system is capable of translating between more languages than it was possible before.

[0037] In addition, the "Knowledge Base" features dictionaries that are able to learn from past references. This feature improves accuracy and efficiency of both the recognition and the inference processes.

[0038] (Practical application: Example 2) Figure 4 is another example of the practical application of this invention. This corresponds to the Step S3 of the translation process described in Figure 1 above.

[0039] In this example, translation process S3 is divided into the following: The system conducts structure analysis of the input document (S3<sub>1</sub>); Fuzzy-inference discards ambiguities from the input document to yield a computer-decipherable document (S3<sub>2</sub>); Fuzzy-inference referred to above discards ambiguous words (if necessary, it may refer a word to the vocabulary dictionary of the "Knowledge Base" (S3<sub>3</sub>); The sentence thus trimmed (or extracted) is recognised (S3<sub>4</sub>);

[0040] The system then conducts a natural language analysis in terms of syntax (S3<sub>5</sub>, S3<sub>6</sub>); and semantics (S3<sub>7</sub>, S3<sub>8</sub>).

[0041] The document thus recognised in terms of vocabulary, syntax and semantics is then translated by using translation dictionary. If the translation is difficult (S3<sub>9</sub>), the document is returned to the fuzzy recognition process (S3<sub>2</sub>) in order to yield translatable sentences before another attempt for translation (S3<sub>10</sub>).

[0042] While this translation process is running, the translated document is referred to the "Knowledge Base" for an analysis of vocabulary, grammar and syntax to sort the order of translated words (S3<sub>11</sub>).

[0043] The following is a detailed description of the process by which the computer discards ambiguities from the original document to yield decipherable sentences.

[0044] The sample input document reads "Hutatsu no buzoku ga sunde iru shimani, hitori no senkyoshi ga tochaku shimashita (A missionary arrived on an island where two tribes were living)". Firstly, the computer chooses one of the following three conditional Rules that may best help recognising the sentence.

[0045] (a) Condition: "if the sentence is decipherable by the computer", Result: "The computer recognises it".

[0046] (b) Condition: "if the sentence is not decipherable by the computer", Result: "The computer discards part of the sentence".

[0047] (c) Condition: "if the sentence is not decipherable by the computer", Result: "The computer refers the sentence to the Knowledge Base".

[0048] If the sentence is decipherable by the computer, the computer selects Rule (a) above, and recognises it. If the sentence is not recognisable, the computer decipheres it. If the sentence is not decipherable, the computer selects Rule (b) and apply max-min analysis to the sentence. If the sentence's probability of success is high, the computer deems it as decipherable and retain the sentence. Otherwise, the computer discards the sentence. Rule (c) concerns sentences that are not computer-decipherable because they contain phrases or words that are unknown to the computer. The computer refers such phrases and words to the (learning) vocabulary dictionary to conduct vocabulary searches.

[0049] The following example uses well-known [Mamudani(?)] method of fuzzy-inference. In this example, a condition (e,  $\Delta e$ ) is fed to Rules (a), (b) and (c). The results of the inference are  $\Delta U1$ ,  $\Delta U2$  and  $\Delta U3$ .  $\Delta U$  is the centre of gravity yielded by overlaying the three results.

[0050] In this example, the sample sentence above is placed on a matrix as follows:

[0051] [Table 1]

U =	
1	Senkyoshi ga (A missionary)
1	Tochaku shimashita (arrived)
-0.9	, (comma)
0.1	sundeiru (living on)
0.6	shimani (an island)
0.8	hutatsu no buzoku ga (two tribes are)
-0.1	hitoride (alone)
-0.9	. (full stop)

[0052] Figure 5 illustrates how the [Mamudani] method analyses the above matrix. As shown in the figure, one of the "fuzzy-labels", i.e., PB (positive big), PM (positive medium), PS (positive small), ZO (zero) or NS (negative small) are attached to each phrase. These labels are standardised as a set of trapezoids on a base of -6 to 6 (the standardisation allows a common usage among many different fuzzy variables).

[0053] Figure 6 illustrate the above by triangular continuous functions. Figure 7 is a max-min chart which shows the clear areas where phrases are either retained or discarded.

[0054] The shaded areas (from 0.1 to -0.1) in Figure 7 is the area where the computer can not determine whether or not to retain the phrase. Phrases that fall into this area are considered under Rule (c) above and are referred to the "Knowledge Base" (the area is called "Knowledge Base zone".) In this example, the result of this process is "Senkyoshi ga hutatsuno buzokuga sumu shimani tochaku shimasita (A missionary arrived on an island where two tribes were living)."

[0055] The second practical application of this invention as illustrated above allows the computer, upon receiving an ambiguous input document, to discard ambiguous words and phrases by fuzzy-inference to yield a document suitable for translation. The document thus yielded is then put to translation, thus improving accuracy of the translation. It also reduces the works needed by the users to check the translation.

[0056] The translation referred to above may be conducted by an expert system which applies inference rules for the translation of the document into the language of destination.

[0057]

[Benefit of this invention] This invention recognises an input document through vocabulary recognition, syntax recognition and semantic analysis using the respective dictionaries of the "Knowledge Base". The recognised document is then translated, using the inference rules of the "Knowledge Base", into the language of the destination. The advantage of this system is as follows:

[0058] (1) Knowledge-based document recognition and inference allows translations between more languages than before.

[0059] (2) Document recognition and inference improves accuracy of the translation.

[0060] (3) Improved accuracy of translation reduces the necessity of checking work by the user.

[0061] (4) The user may send a document in the destination's language by simply inputting an original document, selecting the destination's language and then conducting simple checking work. Thus the integration of electronic mails system [(and the translation system)] is more advanced than before.

[0062] Furthermore, this invention improves accuracy of translation by discarding ambiguities from input documents by using fuzzy-inference.

[Brief explanation of figures]

[Figure 1] Order of processing electronic mails in the practical implementation of this invention.

[Figure 2] An example of screen display image on the sender's terminal.

- [Figure 3] Translation process of the "practical application of this invention: Example 1" above.
- [Figure 4] Translation process of the "practical application of this invention: Example 2" above.
- [Figure 5] Document decipherment using fuzzy-inference employed by the "Example 2" above.
- [Figure 6] An example of "membership function" used by the "Example 2" above.
- [Figure 7] Areas where the computer determines whether or not to discard the phrase.
- [Figure 8] A conventional translation system using the "transfer method".
- [Figure 9] Three conventional methods of translation.
- [Figure 10] Process of translation by conventional methods.

Figure 1

Practical implementation of this invention

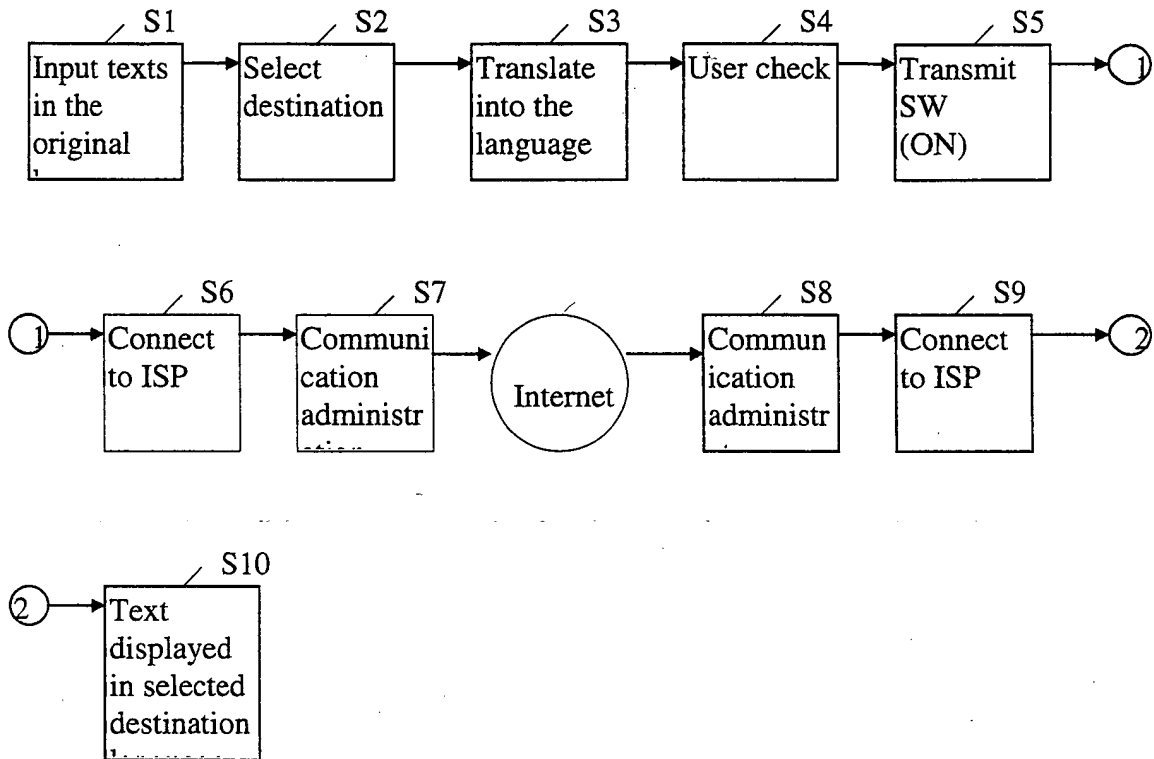


Figure 2

Machine Name		<table border="1"><tr><td>Destination</td><td>1. USA</td><td>▽</td></tr><tr><td></td><td>2. UK</td><td></td></tr><tr><td></td><td>3. Germany</td><td></td></tr><tr><td></td><td>4. France</td><td></td></tr><tr><td></td><td>5. China</td><td></td></tr><tr><td></td><td>6. Korea</td><td></td></tr></table>	Destination	1. USA	▽		2. UK			3. Germany			4. France			5. China			6. Korea	
Destination	1. USA		▽																	
	2. UK																			
	3. Germany																			
	4. France																			
	5. China																			
	6. Korea																			
User Name																				
Password																				
Document Title																				
Transmit	<input type="button" value="ON"/>																			

Figure 4

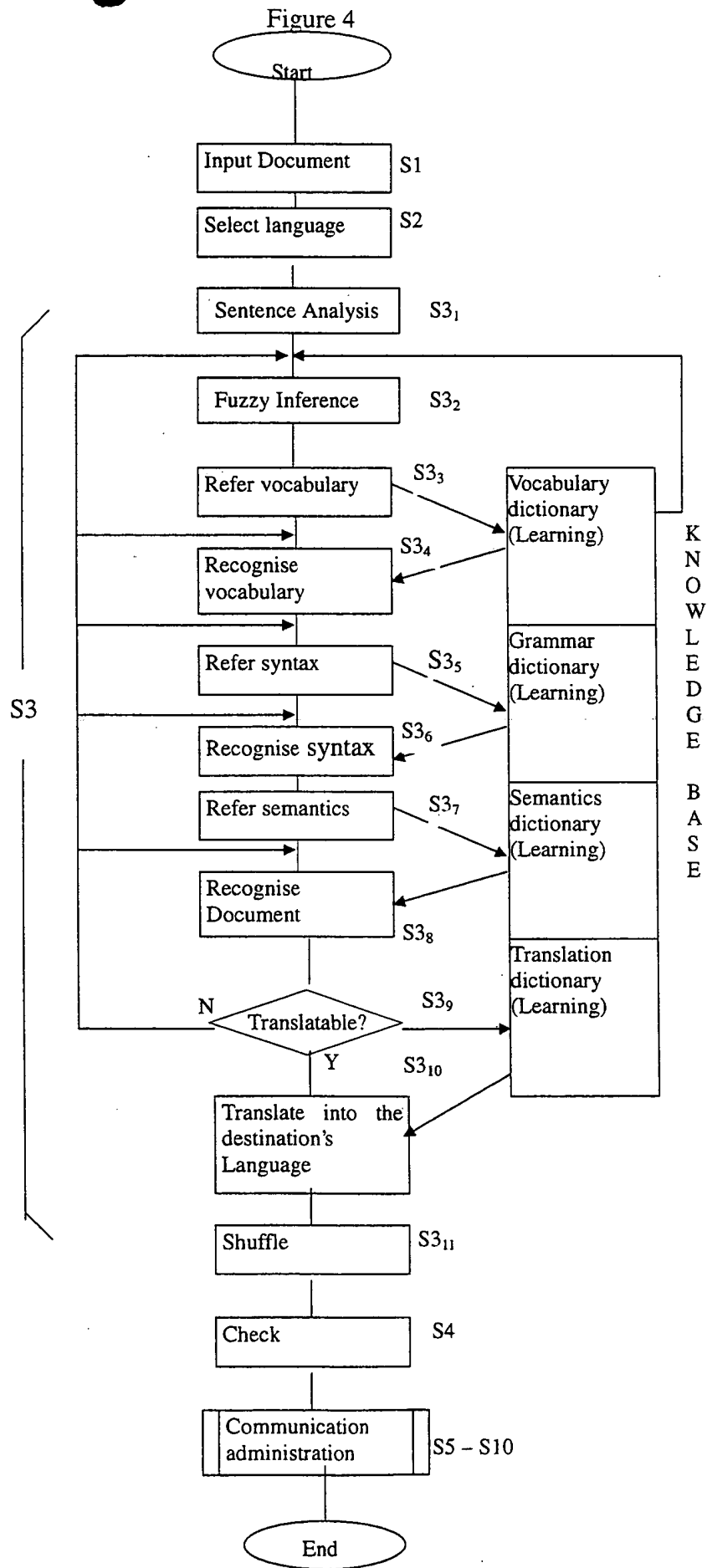


Figure 5

Document Decipherment Using Fuzzy Inference

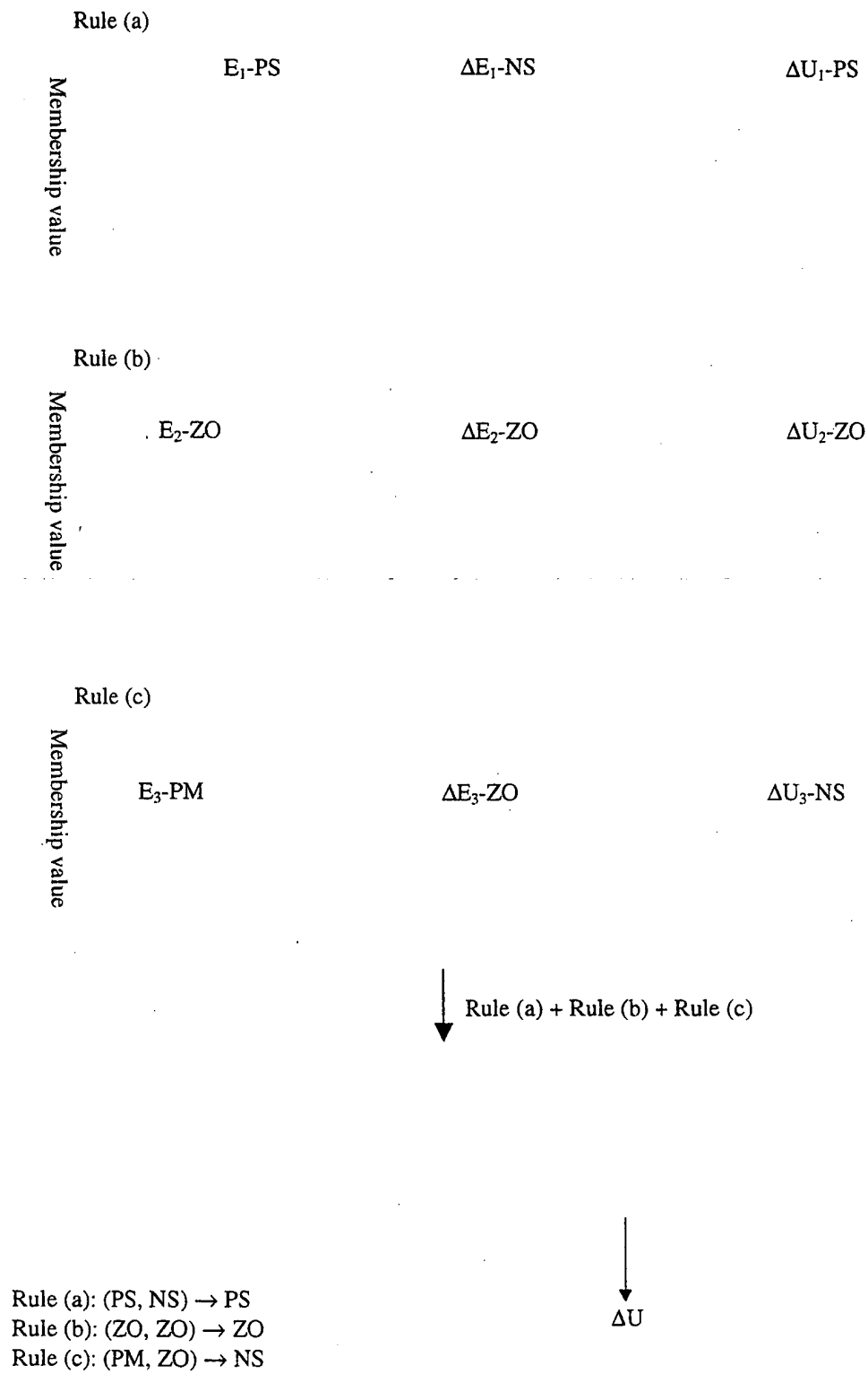




Figure 6

Membership function

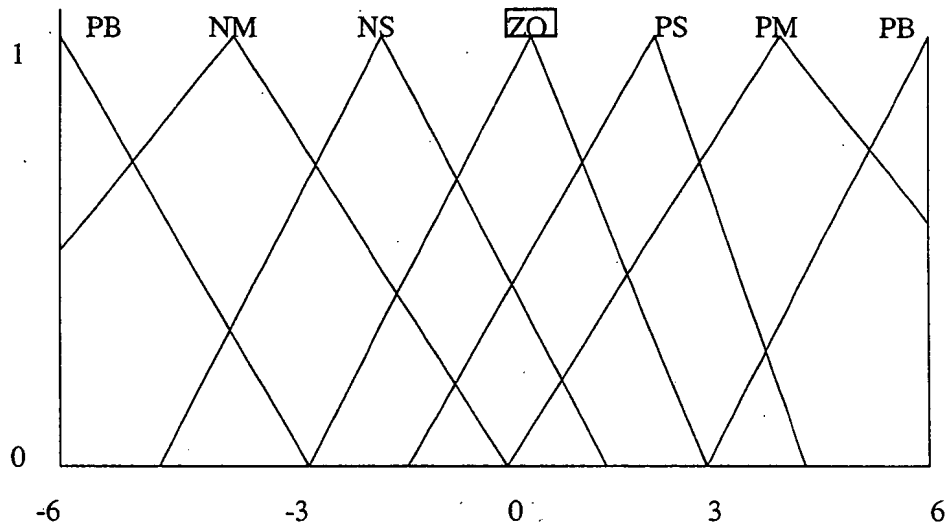


Figure 7

Areas where the computer determines whether or not to discard the phrase.

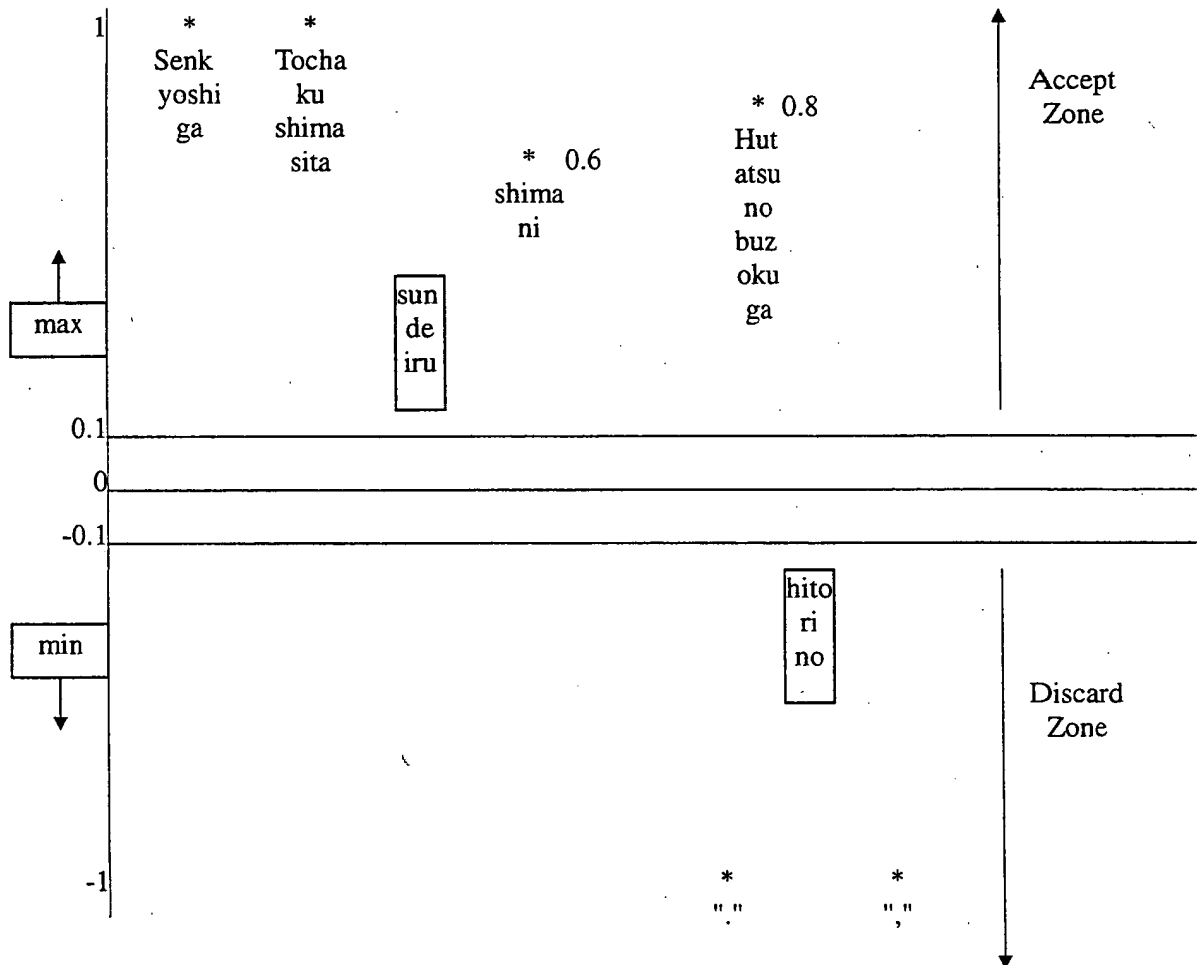


Figure 8

A conventional translation system using the "transfer method".  
(Concept chart)

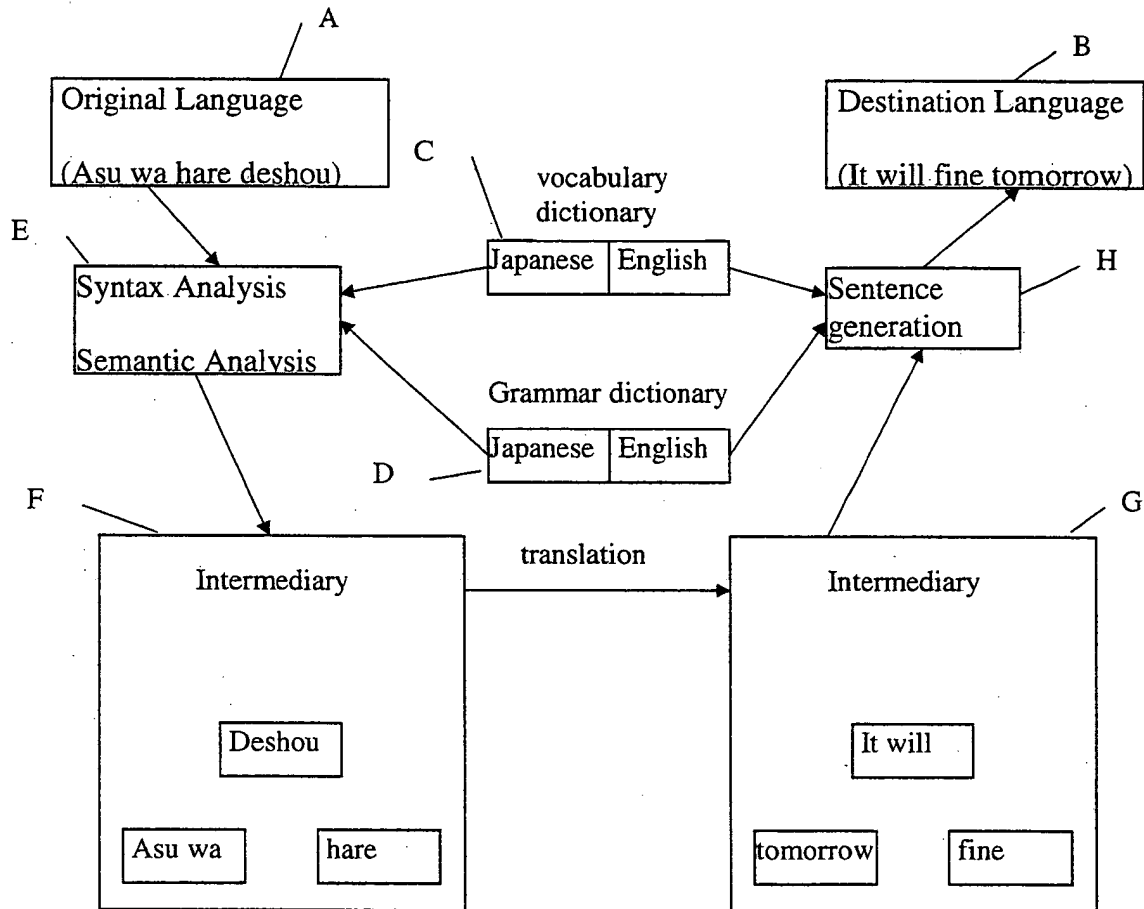


Figure 9

Three conventional methods of translation.

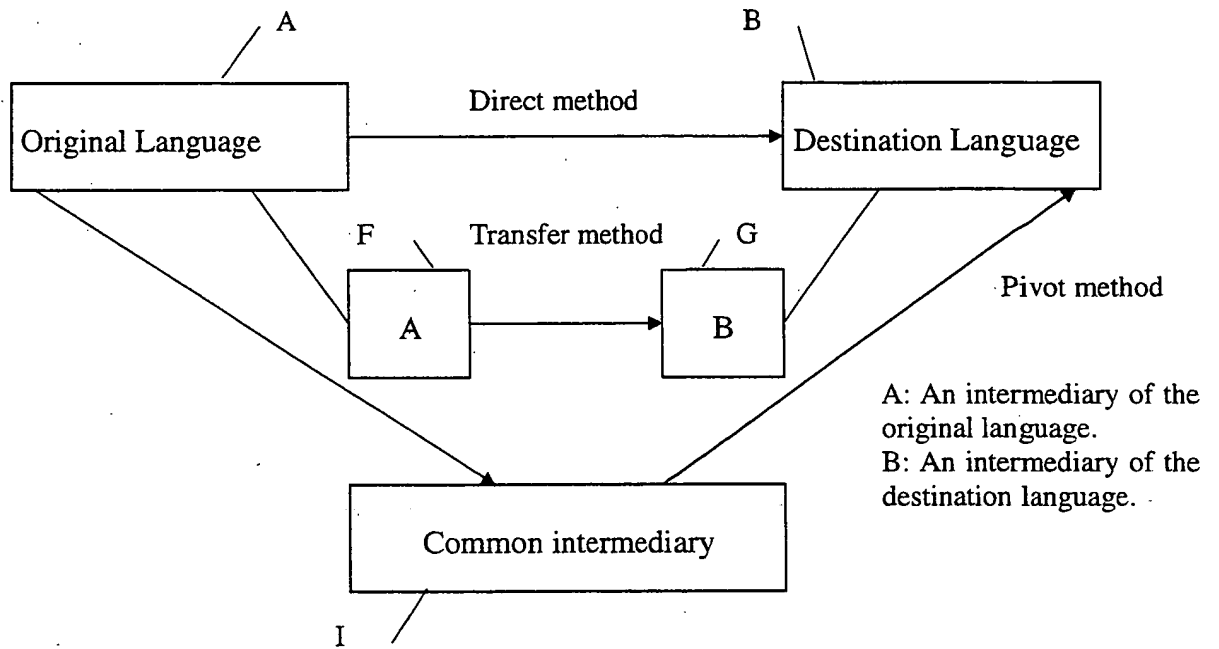


Figure 10

Process of translation by conventional methods

